HW3. Linear Regression

Phuong Anh Trinh - 11200417

October 2022

1 Exercise 1

$$t = y(x, w) + \epsilon$$

Suppose ϵ is Gaussian distributed $N(\mu, \sigma^2)$ and has mean 0
$\Rightarrow \epsilon \sim N(0; \sigma^2)$
$\Rightarrow t = y(x, w) + \epsilon \sim N(y(x, w), \sigma^2)$
$\Rightarrow p(t) = N(t|y(x, w), \sigma^2)$
Consider a data set of inputs $X = \{x_1, x_2, x_3, ..., x_N\}$ with target values $t_1, t_2, t_3, ..., t_N$ (grouped into a column vector denoted by t)

$$p(t|(X, w)).\beta = \prod_{i=1}^{\infty} N(t_n|\frac{y(x, w)}{u}, \frac{\beta^{-1}}{\sigma^2})$$

$$N(t_n|y(x, w), \beta^{-1}) = \frac{1}{\sqrt{2\pi\beta^{-1}}}.e^{\frac{(t - y(x, w)^2)\beta}{2}}$$

It is convenient to take logarithm of the likelihood function

$$\log p(t|X, w, \beta) = \sum_{n=1}^{N} log(N(t_n|y(x_n, w), \beta^{-1})$$

$$= \frac{-\beta}{2} \textstyle\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2$$

$$\max_w \log p(t|X, w, \beta) = -\max_w \frac{-\beta}{2} \sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2$$

$$= \min\nolimits_w \frac{1}{2} \textstyle\sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2$$

We minimize $P = \frac{1}{2} \sum_{n=1}^{N}\{y(x_n, w) - t_n\}^2$ to find w

Suppose

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_n \end{bmatrix}, w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

$\Rightarrow P = \|X_W - t\|_2^2$
$\Rightarrow w = (X^T X)^{-1} X^T t$

2 Exercise 4: Show that $X^T X$ is invertible when X is full rank

When A if full rank, then rank(A) = min$\{m, n\}$, $m \geq n$, then rank(A) = n
$X^T X$ is symmetric, so $X^T X$ is invertible if it is positive definite.

$$\text{If } u^T X^T X u = 0$$
$$\Rightarrow (Xu)^T (Xu) =< Xu, Xu >= \|Xu\|_2^2 = 0$$
$$\Rightarrow Xu = 0 \Rightarrow u = 0$$

Hence $X^T X$ is positive definite. Therefore, $X^T X$ is invertible.